

16K1

1638

1600

RAW SEQUENCE LISTING

DATE: 11/09/2001

PATENT APPLICATION: US/09/445,480A

TIME: 15:48:19

Input Set : A:\Sequence Listing

Output Set: N:\CRF3\11092001\I445480A.raw

SEQUENCE LISTING

4 (1) GENERAL INFORMATION:

6 (i) APPLICANT: Jongsma, Maarten Anthonie
7 Strukelj, Borut
8 Lenarcic, Brigita
9 Gruden, Kristina
10 Turk, Vito
11 Bosch, Hendrik J.
12 Stiekema, Willem Johannes

14 (ii) TITLE OF INVENTION: A Method for Plant Protection Against Insects
15 or Nematodes

17 (iii) NUMBER OF SEQUENCES: 4

19 (iv) CORRESPONDENCE ADDRESS:

20 (A) ADDRESSEE: DLO-Center for Plant Breeding and
21 Reproduction Research

22 (B) STREET: Droevendaalsesteeg 1

23 (C) CITY: Wageningen

24 (D) STATE: N/A

25 (E) COUNTRY: The Netherlands

C--> 26 (F) ZIP: 6708 PB

28 (v) COMPUTER READABLE FORM:

29 (A) MEDIUM TYPE: 3.5 Floppy disk, 1.44 MB

30 (B) COMPUTER: IBM PC compatible

31 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

32 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

34 (vi) CURRENT APPLICATION DATA:

C--> 35 (A) APPLICATION NUMBER: US/09/445,480A

C--> 36 (B) FILING DATE: 07-Jul-2000

37 (C) CLASSIFICATION: C12N15/82

39 (vii) PRIOR APPLICATION DATA:

40 (A) APPLICATION NUMBER: PCT/NL98/00352

41 (B) FILING DATE: June 18, 1998

43 (viii) ATTORNEY/AGENT INFORMATION:

44 (A) NAME: McCLURE, DANIEL R.

45 (B) REGISTRATION NUMBER: 38,962

46 (C) REFERENCE/DOCKET NUMBER: 250308-1020

48 (ix) TELECOMMUNICATION INFORMATION:

49 (A) TELEPHONE: +31 317 477001

50 (B) TELEFAX: +31 317 418094

58 (2) INFORMATION FOR SEQ ID NO: 1:

60 (i) SEQUENCE CHARACTERISTICS:

61 (A) LENGTH: 888 base pairs

62 (B) TYPE: nucleic acid

63 (C) STRANDEDNESS: double

64 (D) TOPOLOGY: linear

66 (ii) MOLECULE TYPE: cDNA

68 (iii) HYPOTHETICAL: NO

RECEIVED
DEC 07 2001
TECH CENTER
ENTERED

RAW SEQUENCE LISTING

DATE: 11/09/2001

PATENT APPLICATION: US/09/445,480A

TIME: 15:48:19

Input Set : A:\Sequence Listing

Output Set: N:\CRF3\11092001\I445480A.raw

RECEIVED
DEC 07 2001
TECH CENTER 1600/2900

```

70      (iv) ANTI-SENSE: NO
72      (vi) ORIGINAL SOURCE:
73          (A) ORGANISM: Actinia equina
75      (ix) FEATURE:
76          (A) NAME/KEY: mat_peptide
77          (B) LOCATION:99..695
79      (ix) FEATURE:
80          (A) NAME/KEY: CDS
81          (B) LOCATION:3..695
83      (ix) FEATURE:
84          (A) NAME/KEY: sig_peptide
85          (B) LOCATION:3..98
87      (ix) FEATURE:
88          (A) NAME/KEY: 5'UTR
89          (B) LOCATION:1..2
91      (ix) FEATURE:
92          (A) NAME/KEY: 3'UTR
93          (B) LOCATION:696..888
95      (x) PUBLICATION INFORMATION:
96          (A) AUTHORS: Gruden, Kristina
97      Strukelj, Borut
98      Popovic, Tatjana
99      Lenarcic, Brigita
100     Bevec, Tadeja
101     Brzin, Joze
102     Kregar, Igor
103     Herzog-Velikonja, Jana
104     Stiekema, Willem J
105     Bosch, Dirk
106          (B) TITLE: The Cysteine Protease Activity of Colorado
107     Potato Beetle (Leptinotarsa decemlineata) Guts,
108     Which is Insensitive to Potato Protease
109     Inhibitors, is Inhibited by Thyroglobulin Type-1
110     Domain Inhibitors
111          (C) JOURNAL: Insect Biochem. Mol. Biol
112          (D) VOLUME: 28
113          (F) PAGES: 549-560
114          (G) DATE: 1998
116     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
118     CT ATG GCT CTT AGC CAA AAC CAA GCC AAG TTT TCC AAA GGA TTC GTC      47
119     Met Ala Leu Ser Gln Asn Gln Ala Lys Phe Ser Lys Gly Phe Val
120     -32      -30      -25      -20
122     GTG ATG ATT TGG GTA CTA TTC ATT GCT TGT GCT ATA ACT TCA ACT GAA      95
123     Val Met Ile Trp Val Leu Phe Ile Ala Cys Ala Ile Thr Ser Thr Glu
124     -15      -10      -5
126     GCT AGT CTA ACC AAA TGC CAA CAG CTC CAG GCC TCG GCT AAC AGT GGT      143
127     Ala Ser Leu Thr Lys Cys Gln Gln Leu Gln Ala Ser Ala Asn Ser Gly
128     1          5          10          15
130     CTG ATA GGT ACT TAT GTA CCA CAA TGC AAA GAA ACG GGA GAG TTC GAA      191

```

RAW SEQUENCE LISTING

DATE: 11/09/2001

PATENT APPLICATION: US/09/445,480A

TIME: 15:48:19

Input Set : A:\Sequence Listing

Output Set: N:\CRF3\11092001\I445480A.raw

```

131 Leu Ile Gly Thr Tyr Val Pro Gln Cys Lys Glu Thr Gly Glu Phe Glu
132                20                25                30
134 GAA AAA CAA TGC TGG GGA TCG ACT GGT TAC TGT TGG TGT GTG GAT GAA      239
135 Glu Lys Gln Cys Trp Gly Ser Thr Gly Tyr Cys Trp Cys Val Asp Glu
136                35                40                45
138 GAT GGA AAA GAG ATT CTA GGA ACC AAG ATC CGT GGA TCT CCG GAT TGC      287
139 Asp Gly Lys Glu Ile Leu Gly Thr Lys Ile Arg Gly Ser Pro Asp Cys
140                50                55                60
142 AGC CGC AGA AAA GCC GCG TTA ACA CTT TGC CAG ATG ATG CAA GCC ATC      335
143 Ser Arg Arg Lys Ala Ala Leu Thr Leu Cys Gln Met Met Gln Ala Ile
144                65                70                75
146 ATT GTT AAT GTC CCT GGT TGG TGT GGC CCT CCA TCG TGT AAA GCT GAC      383
147 Ile Val Asn Val Pro Gly Trp Cys Gly Pro Pro Ser Cys Lys Ala Asp
148    80                85                90                95
150 GGC AGT TTT GAC GAG GTT CAG TGC TGC GCA AGT AAT GGA GAA TGC TAC      431
151 Gly Ser Phe Asp Glu Val Gln Cys Cys Ala Ser Asn Gly Glu Cys Tyr
152                100                105                110
154 TGT GTG GAT AAG AAA GGA AAA GAA CTT GAA GGC ACA AGA CAA CAG GGA      479
155 Cys Val Asp Lys Lys Gly Lys Glu Leu Glu Gly Thr Arg Gln Gln Gly
156                115                120                125
158 AGG CCA ACC TGC GAA AGA CAC CTA AGC GAA TGC GAG GAA GCT CGA ATC      527
159 Arg Pro Thr Cys Glu Arg His Leu Ser Glu Cys Glu Glu Ala Arg Ile
160                130                135                140
162 AAG GCG CAT TCA AAC AGT CTT CGT GTT GAG ATG TTC GTG CCA GAG TGT      575
163 Lys Ala His Ser Asn Ser Leu Arg Val Glu Met Phe Val Pro Glu Cys
164                145                150                155
166 TTA GAA GAT GGA TCA TAT AAC CCA GTA CAG TGC TGG CCT AGC ACA GGA      623
167 Leu Glu Asp Gly Ser Tyr Asn Pro Val Gln Cys Trp Pro Ser Thr Gly
168    160                165                170                175
170 TAC TGT TGG TGC GTC GAT GAA GGA GGG GTA AAG GTA CCA GGT TCC GAT      671
171 Tyr Cys Trp Cys Val Asp Glu Gly Gly Val Lys Val Pro Gly Ser Asp
172                180                185                190
174 GTC AGA TTT AAA CGC CCC ACA TGC TAAGAAAAAC ACAGTGAACA AAGTGGCTAG      725
175 Val Arg Phe Lys Arg Pro Thr Cys
176                195
178 TTCCAGATC GAAAAATAACT ACAAAGGATT AATAAAATGT TAAAAATAATT TCTCAATTCG      785
180 GCTGTGATAT ATTTTTCCTA AGATAATTGA ATCTGCATGT AGTTAACAGA AAACAATCTC      845
182 AACTAGAAAT AAAGACTACG GTAATAATGA CAAAAAATAA AAA                      888
185 (2) INFORMATION FOR SEQ ID NO: 2:
187     (i) SEQUENCE CHARACTERISTICS:
188         (A) LENGTH: 231 amino acids
189         (B) TYPE: amino acid
190         (D) TOPOLOGY: linear
192     (ii) MOLECULE TYPE: protein
193     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
195 Met Ala Leu Ser Gln Asn Gln Ala Lys Phe Ser Lys Gly Phe Val Val
196 -32    -30                -25                -20
198 Met Ile Trp Val Leu Phe Ile Ala Cys Ala Ile Thr Ser Thr Glu Ala
199    -15                -10                -5

```

RAW SEQUENCE LISTING

DATE: 11/09/2001

PATENT APPLICATION: US/09/445,480A

TIME: 15:48:19

Input Set : A:\Sequence Listing

Output Set: N:\CRF3\11092001\I445480A.raw

```

201 Ser Leu Thr Lys Cys Gln Gln Leu Gln Ala Ser Ala Asn Ser Gly Leu
202 1 5 10 15
204 Ile Gly Thr Tyr Val Pro Gln Cys Lys Glu Thr Gly Glu Phe Glu Glu
205 20 25 30
207 Lys Gln Cys Trp Gly Ser Thr Gly Tyr Cys Trp Cys Val Asp Glu Asp
208 35 40 45
210 Gly Lys Glu Ile Leu Gly Thr Lys Ile Arg Gly Ser Pro Asp Cys Ser
211 50 55 60
213 Arg Arg Lys Ala Ala Leu Thr Leu Cys Gln Met Met Gln Ala Ile Ile
214 65 70 75 80
216 Val Asn Val Pro Gly Trp Cys Gly Pro Pro Ser Cys Lys Ala Asp Gly
217 85 90 95
219 Ser Phe Asp Glu Val Gln Cys Cys Ala Ser Asn Gly Glu Cys Tyr Cys
220 100 105 110
222 Val Asp Lys Lys Gly Lys Glu Leu Glu Gly Thr Arg Gln Gln Gly Arg
223 115 120 125
225 Pro Thr Cys Glu Arg His Leu Ser Glu Cys Glu Glu Ala Arg Ile Lys
226 130 135 140
228 Ala His Ser Asn Ser Leu Arg Val Glu Met Phe Val Pro Glu Cys Leu
229 145 150 155 160
231 Glu Asp Gly Ser Tyr Asn Pro Val Gln Cys Trp Pro Ser Thr Gly Tyr
232 165 170 175
234 Cys Trp Cys Val Asp Glu Gly Gly Val Lys Val Pro Gly Ser Asp Val
235 180 185 190
237 Arg Phe Lys Arg Pro Thr Cys
238 195

```

240 (2) INFORMATION FOR SEQ ID NO: 3:

242 (i) SEQUENCE CHARACTERISTICS:

243 (A) LENGTH: 696 base pairs

244 (B) TYPE: nucleic acid

245 (C) STRANDEDNESS: double

246 (D) TOPOLOGY: linear

248 (ii) MOLECULE TYPE: cDNA

250 (iii) HYPOTHETICAL: NO

252 (iv) ANTI-SENSE: NO

254 (vi) ORIGINAL SOURCE:

255 (A) ORGANISM: Actinia equina

257 (vii) IMMEDIATE SOURCE:

258 (B) CLONE: optimized gene for expression in plants

260 (ix) FEATURE:

261 (A) NAME/KEY: CDS

262 (B) LOCATION:1..693

264 (ix) FEATURE:

265 (A) NAME/KEY: mat_peptide

266 (B) LOCATION:97..693

268 (ix) FEATURE:

269 (A) NAME/KEY: sig_peptide

270 (B) LOCATION:1..693

273 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

RAW SEQUENCE LISTING

DATE: 11/09/2001

PATENT APPLICATION: US/09/445,480A

TIME: 15:48:19

Input Set : A:\Sequence Listing

Output Set: N:\CRF3\11092001\I445480A.raw

```

275 ATG GCT CTT AGC CAG AAC CAG GCC AAG TTT TCC AAG GGA TTC GTC GTG      48
276 Met Ala Leu Ser Gln Asn Gln Ala Lys Phe Ser Lys Gly Phe Val Val
277 -32 -30 -25 -20
279 ATG ATT TGG GTA CTA TTC ATT GCT TGT GCT ATC ACT TCA ACT GAA GCT      96
280 Met Ile Trp Val Leu Phe Ile Ala Cys Ala Ile Thr Ser Thr Glu Ala
281 -15 -10 -5
283 AGT CTA ACG AAA TGC CAA CAG CTG CAG GCC TCG GCT AAC AGT GGT CTG      144
284 Ser Leu Thr Lys Cys Gln Gln Leu Gln Ala Ser Ala Asn Ser Gly Leu
285 1 5 10 15
287 ATA GGT ACT TAT GTA CCA CAA TGC AAA GAA ACT GGA GAG TTT GAA GAA      192
288 Ile Gly Thr Tyr Val Pro Gln Cys Lys Glu Thr Gly Glu Phe Glu Glu
289 20 25 30
291 AAG CAA TGC TGG GGA TCG ACT GGT TAC TGT TGG TGT GTG GAT GAA GAT      240
292 Lys Gln Cys Trp Gly Ser Thr Gly Tyr Cys Trp Cys Val Asp Glu Asp
293 35 40 45
295 GGA AAA GAG ATT CTA GGT ACA AAG ATC CGT GGA TCT CCA GAC TGC AGT      288
296 Gly Lys Glu Ile Leu Gly Thr Lys Ile Arg Gly Ser Pro Asp Cys Ser
297 50 55 60
299 CGC AGA AAA GCT GCC TTA ACA CTT TGC CAG ATG ATG CAA GCC ATC ATT      336
300 Arg Arg Lys Ala Ala Leu Thr Leu Cys Gln Met Met Gln Ala Ile Ile
301 65 70 75 80
303 GTG AAT GTC CCT GGT TGG TGT GGA CCT CCA TCA TGT AAA GCT GAC GGC      384
304 Val Asn Val Pro Gly Trp Cys Gly Pro Pro Ser Cys Lys Ala Asp Gly
305 85 90 95
307 AGT TTT GAC GAG GTT CAG TGC TGC GCA AGT AAT GGA GAA TGC TAC TGT      432
308 Ser Phe Asp Glu Val Gln Cys Cys Ala Ser Asn Gly Glu Cys Tyr Cys
309 100 105 110
311 GTG GAT AAG AAA GGA AAA GAA CTT GAA GGC ACA AGA CAA CAG GGA AGG      480
312 Val Asp Lys Lys Gly Lys Glu Leu Glu Gly Thr Arg Gln Gln Gly Arg
313 115 120 125
315 CCA ACC TGC GAA AGA CAC CTA AGC GAA TGC GAG GAG GCT CGT ATC AAG      528
316 Pro Thr Cys Glu Arg His Leu Ser Glu Cys Glu Glu Ala Arg Ile Lys
317 130 135 140
319 GCA CAT TCA AAC AGT CTT CGT GTT GAG ATG TTC GTG CCA GAG TGT TTA      576
320 Ala His Ser Asn Ser Leu Arg Val Glu Met Phe Val Pro Glu Cys Leu
321 145 150 155 160
323 GAA GAT GGA TCT TAC AAC CCT GTA CAG TGC TGG CCT AGC ACA GGA TAC      624
324 Glu Asp Gly Ser Tyr Asn Pro Val Gln Cys Trp Pro Ser Thr Gly Tyr
325 165 170 175
327 TGT TGG TGC GTC GAT GAA GGA GGG GTA AAG GTT CCA GGT TCC GAC GTC      672
328 Cys Trp Cys Val Asp Glu Gly Gly Val Lys Val Pro Gly Ser Asp Val
329 180 185 190
331 AGA TTC AAA CGT CCC ACA TGC TAA      696
332 Arg Phe Lys Arg Pro Thr Cys
333 195
336 (2) INFORMATION FOR SEQ ID NO: 4:
338 (i) SEQUENCE CHARACTERISTICS:
339 (A) LENGTH: 231 amino acids
340 (B) TYPE: amino acid

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/445,480A

DATE: 11/09/2001

TIME: 15:48:20

Input Set : A:\Sequence Listing

Output Set: N:\CRF3\11092001\I445480A.raw

L:26 M:220 C: Keyword misspelled or invalid format, [(F) ZIP:]
L:35 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:36 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]